



INSTITUTE FOR DEFENSE ANALYSES

The Accuracy of Independent Estimates of the Procurement Costs of Major Systems

David L. McNicol, Project Leader

Karen W. Tyson

John R. Hiller

Harley A. Cloud

Joshua A. Minix

August 2005

Approved for public release;
distribution unlimited.

IDA Paper P-3989

Log: H 05-000207

20060216 016

This work was conducted under contract DASW01 04 C 0003, Task AO-7-2454, for the Office of the Director, Acquisition Resources and Analysis. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

© 2005, 2006 Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882 • (703) 845-2000.

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227-7013 (NOV 95).

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-3989

**The Accuracy of Independent Estimates of
the Procurement Costs of Major Systems**

David L. McNicol, Project Leader
Karen W. Tyson
John R. Hiller
Harley A. Cloud
Joshua A. Minix

PREFACE

The Institute for Defense Analyses (IDA) prepared this paper for the Office of the Director, Acquisition Resources and Analysis, under a task titled "Rolling Capture of Acquisition Lessons Learned." This paper explores the accuracy of independent estimates of procurement cost for a sample of 25 major weapon programs. It partially fulfills the task objective of providing periodic studies that draw acquisition lessons from experience on sets of major programs.

David R. Graham and J. Richard Nelson of IDA were the technical reviewers for this paper.

TABLE OF CONTENTS

Summary	S-1
I. Introduction	1
II. Methodology	5
A. Distribution of Procurement Cost Growth.....	5
B. Metrics	7
C. Self-Fulfilling Prophecies	11
D. Concluding Comments.....	14
III. Data	15
A. SCP and ICE	15
B. What Was the Basis of the Funding Decision?.....	17
C. Procurement Cost Growth Defined.....	18
D. Causes of Cost Growth	19
E. Concluding Comments.....	20
IV. Results	23
A. Frequency of Failures to Identify an Unrealistically Optimistic SCP	23
B. Frequency of False Alarms	24
C. Self-Fulfilling Prophecy?.....	26
D. Concluding Comments.....	26
V. Recommendations	29
A. Conducting Further Studies	29
B. Establishing a Process to Track Cost Issues	31
C. Capturing Key Funding Data	32
D. Concluding Comments.....	33
Appendix A: Deflation, Quantity Normalization, and Adjustment for Content Changes	A-1
Appendix B: Systems Included in Sample.....	B-1
Abbreviations.....	C-1

LIST OF FIGURES

1. Distribution of Adjusted Procurement Cost Growth from the Milestone II Baseline for 138 Major Weapon System Programs	6
2. Histogram of the Percentage Difference between the Independent Cost Estimate for Procurement and the Service Cost Position for Procurement for 25 Major Defense Acquisition Programs.....	16
3. Distribution of Adjusted Procurement Cost Growth from the Milestone II Baseline for 21 Major Defense Acquisition Programs.....	19

LIST OF TABLES

1.	Average Growth in Procurement Cost Estimates from Milestone II Before and After the Inauguration of Independent Costing in the Defense Department	2
2.	Principal Markers of Realistically and Unrealistically Funded Procurement Programs.....	13
3.	Assessed Accuracy of the ICEs for 17 MDAPs for Which Procurement Cost Was Not Raised as an Issue at the DAB Milestone II Review	24
4.	Assessed Accuracy of the ICEs for 8 MDAPs for Which Procurement Cost Was Raised as an Issue at the DAB Milestone II Review	26

SUMMARY

The Milestone Decision Authority (MDA) is required by statute to consider an independent cost estimate before permitting a major defense acquisition program (MDAP) to enter the Engineering and Manufacturing Development (EMD) phase (Milestone II). The MDA must again consider an Independent Cost Estimate (ICE) before permitting production and deployment of the system. When decision authority for the program is retained in the Office of the Secretary of Defense (OSD), the statute defines an "independent" estimate as one made by an organization "that is not under the supervision, direction, or control of the military department...directly responsible for carrying out the development or acquisition of the program."¹

The Institute for Defense Analyses assessed the accuracy of ICEs made to support milestone reviews of major programs in the Department of Defense (DoD) acquisition process. Do these estimates usually overestimate what prove to be the actual costs of the program proposed for milestone funding authority? Do they, instead, systematically understate costs? Or has the accuracy of the independent estimates in most cases proved to be acceptable to decision makers?

Our initial sample included 63 systems that passed Milestone (MS) II between 1985 and 1998, inclusive. For only 25 of these were we able to locate the ICE, the Service Cost Position (SCP), and the other data required for our analysis. For these 25 systems, we compared the ICE and the SCP with observed growth in procurement costs, adjusted for inflation, changes in the total quantity acquired, and cost growth due to unforced additions of program content made after Milestone (MS) II (usually enhancements in performance).

The procurement estimates whose accuracy we reviewed were made at the time that the programs were entering the EMD phase. At that point, many of the large features of the program that determine procurement cost had been selected. Development had not

¹ 10 U.S.C. §2434, paragraph (b), subparagraph (1)A(i). The EMD phase of the acquisition process is now referred to as System Development and Demonstration (SDD), and 10 USC 2434 has been amended to reflect this change. We use the older terminology because that is what is used in the sources of our data.

been completed; however, procurement expenditures were still several years in the future, and, once begun, would continue for upwards of 10 years. Consequently, it is not unreasonable to expect a considerable degree of uncertainty in the estimates made at MS II.

We defined the Acceptable range of growth in procurement cost as -25 percent to +25 percent. If observed cost growth exceeded 25 percent, we classified the MS II ICE as Too Low, and if it was less than -25 percent, we classified the MS II estimate as Too High.²

The lower bound of our Acceptable range is arbitrary, but the upper bound reflects statute. In 1982, the Congress established requirements for reporting the cost growth experienced on individual MDAPs.³ The Congress must be notified whenever growth in the unit procurement cost (as defined in the statute) is 15 percent or more against the established program baseline. If unit procurement cost growth exceeds 25 percent, the Under Secretary of Defense (Acquisition, Technology and Logistics) must notify the Congress and certify that the acquisition program is essential to national security and that the revised cost estimate is reasonable. The statute further precludes obligating program funds for any of the program's major contracts if the required certifications are not provided to the Congress. Thus, cost growth of 25 percent can reasonably be taken as an indication of how much procurement cost growth is "too much" from the congressional point of view.

We assigned the ratings to the ICEs separately for two subsamples: the MDAPs for which procurement cost probably was not raised as an issue at the Defense Acquisition Board (DAB) and those for which it probably was. In the first of these cases, the principal concern is that the ICE failed to identify SCPs that proved to be overly optimistic. The main concern in the second case is that the ICE proves to be a false alarm; that is, it indicates that the SCP was unrealistically low when, in fact, it was not.

² This statement assumes that cost growth is stated relative to the ICE. It is important to note this because cost growth reported in the MS II Selected Acquisition Reports (SARs) is relative to the funding profile established at the time of the SAR, and this often is not based on the ICE.

³ Pub. L. 97-252 (Defense Authorization Act for Fiscal Year 1983), Title XI, Sec. 1107(a) (1), currently codified as 10 USC 2433. The amendment that led to this provision was introduced by Senator Sam Nunn and Congressman Dave McCurdy. The events that trigger these reporting requirements are often referred to as Nunn-McCurdy breaches.

There were 17 MDAPs in the usable sample for which procurement cost probably was not an issue at the MS II DAB for the system. This ordinarily was the case if the MS II procurement ICE was within 10 percent of the procurement SCP. In these cases, the DAB effectively took the rough agreement of the ICE with the SCP to indicate that the SCP was realistic. For that reason, for these MDAPs we assigned the ICE to the same category as the SCP.

Table S-1 presents our results for this set of cases. In 80 percent (14 of 17) of the MDAPs for which the ICE was roughly equal to the SCP, the SCP proved to be a realistic estimate by the criteria of this study. In two cases the ICE failed to identify an SCP that proved to be unrealistically optimistic, and in one case the ICE failed to identify an SCP that proved to be unrealistically high.

Table S-1. Assessed Accuracy of the ICEs for 17 MDAPs for Which Procurement Cost Was Not Raised as an Issue at the DAB Milestone II Review

	Too Low	Acceptable	Too High	Total
SCP	2	14	1	17
ICE	2	14	1	17

The ICE exceeded the SCP by more than 10 percent for 8 of the 25 systems in the sample. Cost, or aspects of the program closely tied to cost, probably was an issue at the MS II DAB in each of these cases. Table S-2 summarizes our assessments for these 8. The ICE was a false alarm in 25 percent of these cases (2 of 8). The ICE correctly identified an SCP that was unrealistically low in the other 75 percent of the cases (6 of 8). The weight placed on this assessment, however, must be tempered by the fact that two programs for which the ICE was scored as Acceptable used unprecedented contracting strategies, one was cancelled while in EMD, and a fourth was restructured after an initial unsuccessful attempt to gain MS II approval.

Table S-2. Assessed Accuracy of the ICEs for 8 MDAPs for Which Procurement Cost Was Raised as an Issue at the DAB Milestone II Review

	Too Low	Acceptable	Too High	Total
SCP	6	1	1	8
ICE	2	4	2	8

The ICE was used as the basis for procurement funding as a whole for only one of the 25 MDAPs in our sample. It is necessary for this case to consider the possibility that

funding the ICE was a self-fulfilling prophecy. We did so by looking for the unequivocal signs that a program was overfunded:

- Addition of substantial program content beyond that approved at MS II without addition of funding; and
- Removal of funding without any reduction of program content.

We found no evidence that either of these had occurred. Moreover, adjusting for inflation, total quantity procured, and content changes, the MDAP in question overran its MS II procurement baseline—which reflected the ICE—by 20 percent to 25 percent. We concluded that there is no evidence in this case that funding to the ICE was a self-fulfilling prophecy.

Our results on the accuracy of the 25 ICEs we had sufficient data to score can be summarized as follows:

- The sample included 17 cases in which procurement cost was not raised as an issue at the DAB; the ICE failed to identify the 3 of these that proved to be unrealistically low or high.
- The ICE was a false alarm in two of the 8 cases in which cost was raised at the DAB.

These conclusions provide a rough guide on the reliability of ICEs for use by senior DoD officials involved in the acquisition of major weapon systems.

This study offers nine recommendations directed towards providing further insights into the accuracy of ICEs and SCPs. The first five of these are topics for further study on the accuracy of the ICEs and SCPs. The five possible studies identified would all be limited in what they could achieve by availability of data. There is no doubt, however, that each could be undertaken, and there is a reasonable prospect that each could shed some further light on the accuracy of ICEs.

The other four recommendations are of a different character, and most would be controversial to some degree. These fall into two groups. First, the study points to two processes for tracking cost issues in ongoing MDAPs:

- Maintain validated cost, schedule, and performance tracks on all MDAPS.
- Track significant cost issues that reached the DAB in ongoing MDAPs.

The first of these would require a large effort. The other would require less effort, but its purpose obviously is as much quality assurance as analytical, and hence might raise potentially difficult issues of organizational authority and accountability.

The remaining set of recommendations is concerned with effectively capturing the data on funding required to assess the accuracy of the ICEs and SCPs:

- Record the basis of Milestone B funding decisions.
- Include the quantity and funding levels specified in MDAP Acquisition Program Baselines as mandatory guidance in the Joint Programming Guidance.

Recording clearly the basis of the Milestone funding decisions would not require a great deal of additional effort, and we do not see any reason why it would encounter strong opposition. Implementing the final recommendation would essentially be a matter of specifying in the relevant DoD regulations that USD(AT&L) send the APBs to the Director of PA&E, and that the Director of PA&E ensure they are included as mandatory guidance in the Joint Planning Guidance. In fact, this change arguably simply codifies current guidance. It would, however, require approval (through a Program Change Proposal or Budget Change Proposal) of changes that the Services now often make themselves in formulating their Budget Estimate Submissions and Program Objective Memoranda. This recommendation has consequences that go considerably beyond efficient capture of funding data required to evaluate the accuracy of the ICEs and SCPs, and we expect it would be controversial.

I. INTRODUCTION

The Milestone Decision Authority (MDA) is required by statute to consider an independent cost estimate before permitting a major defense acquisition program (MDAP) to enter the Engineering and Manufacturing Development (EMD) phase.¹ The MDA must again consider an independent cost estimate before permitting production and deployment of the system. For those programs for which milestone decision authority is retained in the Office of the Secretary of Defense (OSD), the statute defines an “independent” cost estimate as one made by an organization “that is not under the supervision, direction, or control of the military department, Defense Agency, or other component of the Department of Defense that is directly responsible for carrying out the development or acquisition of the program....”²

The Under Secretary of Defense (Acquisition, Technology and Logistics) (USD[AT&L]) ordinarily is the MDA for MDAPs that have not been delegated to one of the military departments. Since 1993, the independent estimates required by USD(AT&L) in this capacity have been made by the OSD Cost Analysis Improvement Group (CAIG). From the inauguration of independent costing in the Department of Defense (DoD) in 1972 until about 1993, the responsibility for making the independent estimates effectively was shared by the CAIG and cost offices located in the secretariats of the Departments of the Army, Navy, and the Air Force.³

The independence required by the statute is a means, not an end in itself. The intent of the statute, and the DoD regulations that implement it, clearly is that the MDA be presented with a realistic estimate of the cost of the program he is asked to approve.⁴ This

¹ 10 USC 2434. This phase of the acquisition process is now referred to as System Development and Demonstration (SDD), and 10 USC 2434 has been amended to reflect this change. We use the older terminology because that is what is used in the sources of our data.

² 10 USC 2434 paragraph (b) subparagraph (1)A(i).

³ The statutory requirement for an independent estimate was imposed in 1983, about a decade after the CAIG was established.

⁴ See Donald Srull, ed. *The Cost Analysis Improvement Group: A History*, Logistics Management Institute, 1998; and David L. McNicol, “Cost Growth in Major Weapon Procurement Programs,” Institute for Defense Analyses, Paper P-3832 (Nonstandard), 2004, especially pp. 6–10.

paper reports the results of an attempt to evaluate whether independent estimates of procurement cost have in fact been realistic estimates of the costs of programs as they come up for approval to begin EMD. Do they usually overestimate what prove to be the actual procurement costs of the program proposed for milestone funding authority? Do they, instead, systematically understate costs? Or has the accuracy of the independent estimates been acceptable to decision makers in most cases?

Several studies have attempted to infer something about the effectiveness of the Defense Department's independent weapon system costing from trends over time in procurement cost growth. Table 1 presents data along these lines from a recent IDA study.

Table 1. Average Growth in Procurement Cost Estimates from Milestone II Before and After the Inauguration of Independent Costing in the Defense Department

	Pre-CAIG ^a	1974-83	1984-92	1993-97
Average Cost Growth ^b	30%	16%	16%	18%
Number of Systems	15	40	57	26
Number with $\geq 30\%$ Cost Growth	6 (40%)	8 (20%)	11 (19%)	10 (38%)

Source: Abstracted from Table 8, p. 47, of David L. McNicol, "Cost Growth in Major Weapon Procurement Programs," Institute for Defense Analyses, Paper P-3832 (Nonstandard), 2004.

^a FY 1970-73. The CAIG was established in mid-FY 1972, but FY 1974 was the first full year it was able to influence estimates for milestone reviews.

^b PA&E data extracted from Selected Acquisition Reports. Average growth in procurement cost in constant dollars, normalized to Milestone II inventory objective and partially normalized for exogenous changes in program content.

This table shows the average growth⁵ in procurement costs from the Milestone (MS) II baseline for major weapon acquisition programs for four sub-periods: pre-CAIG and three subsequent sub-periods, separated at the years in which the Defense Department's independent costing process was strengthened.⁶ During the decade after the

⁵ Would it have been better to use an average in which cost growth was weighted by program size? Such a weighted average clearly is appropriate if the underlying issue is the effect of cost growth on the budget, since a given percentage costing error is more important for a large program than a small program. We do not believe that the use of a weighted average is called for in the present context, however. Use of a weighted average would imply that the decision maker values accuracy in the cost estimate for a large MDAP more than he does for a small MDAP. It is our understanding that this is not the case. Rather, in the Defense Acquisition Board context, the USD(AT&L) generally looks to cost estimators to provide estimates of a uniform quality across all sizes of MDAPS.

⁶ Consideration of an independent cost estimate at MS II and again at the start of production became a statutory requirement in 1983. In 1992, the CAIG staff was substantially increased, and the CAIG assumed full responsibility for the provision of the independent cost estimates required by statute.

introduction of independent costing, average procurement cost growth was a little more than half what it had been during the preceding 5 pre-CAIG years. Note also that the fraction of systems with procurement cost growth of at least 30 percent was about half that of the pre-CAIG period. Substantially the same is true for the data for 1984–1992. The situation for FY 1993–1997 is a little more complicated. The fraction of systems with procurement cost growth of at least 30 percent returned to substantially its level before the introduction of independent costing. The average cost growth of these high cost-growth systems, however, was only about 60 percent of the pre-CAIG level.⁷

Evidence such as that in Table 1 casts independent estimates of the costs of major weapon procurement programs in a favorable light.⁸ It does not, however, directly evaluate the accuracy of those estimates, and hence is of limited value to a decision maker interested in knowing how reliable the independent estimates tend to be. In particular, evidence like that in Table 1 leaves open the possibility that the post-CAIG reduction in cost growth was more the result of the process established to review Service cost estimates than it was of the independent estimates as such.

In 1992 the DoD Inspector General released a report on the independent costing process in the Defense Department. This report largely was concerned with establishing the extent to which the actual independent costing process followed the guidelines laid down in the relevant DoD documents. It did not go into the accuracy of the independent cost estimates.

To our knowledge, the present study is the first attempt to assess directly the accuracy of a sample of independent cost estimates. This is remarkable. In the more than three decades they have been used, independent estimates have had a significant role in important decisions about many major programs, and they have helped to establish the credibility of the MDAP budget requests for major procurement programs the Defense

⁷ This was a period of intense concern with acquisition reform, and policy on realistic costing was more relaxed. This relaxation was a matter of the stringency with which the policy was enforced, however, not of its abandonment. For more on this point, see McNicol, "Cost Growth in Major Weapon Procurement Programs," pp. 53–54.

⁸ Other studies have also found that average cost growth was noticeably lower in MDAPs that passed MS II after the introduction of independent costing. References to early studies that make this point can be found in Congressional Budget Office, "Cost Growth in Weapon Systems: Some Recent Experience and Possible Remedies," report prepared for the Senate Committee on Governmental Affairs, Washington, DC: Government Printing Office, 1983.

Department submits to the Congress. Yet we lack any published evaluations of how accurate those estimates are.

Information presented in the course of this study suggests that the lack of previous analyses may be due mainly to the difficulty of acquiring the necessary data. This perhaps seems surprising since establishing the accuracy of independent estimates of procurement cost simply requires comparing the forecast of cost—that is, the MS II estimate—with the procurement costs actually incurred. There are, however, two categories of difficulties in executing this simple study design. First, on closer inspection the apparently straightforward task of comparing the independent estimate of procurement cost with the actual procurement cost does present methodological problems. These have to do with appropriate recognition of the context in which the independent estimates are used, and whether basing a procurement funding decision on the independent estimate is a self-fulfilling prophecy. Chapter II identifies these problems specifically and outlines how we deal with them in this report.

The second and much less easily resolved set of problems we encountered is that of capturing the data necessary to make apples-to-apples comparisons of MS II procurement cost estimates with realized costs. The data we used and the main shortcomings of them are described in Chapter III.

We report in Chapter IV the results we obtain on the accuracy of independent procurement cost estimates. Our goals in this report were, first, to lay out succinctly but completely the various steps that must go into an assessment of the accuracy of independent estimates of the costs of weapon procurement programs, and then to get the best estimates we could within the study's limitations of time and resources. The data problems we encountered were intractable, and we could not resolve all of them in a satisfactory manner. Consequently, the results reported in Chapter IV are indicative, not definitive.

Our recommendations are provided in Chapter V. These concern the sorts of follow-on analyses that appear to be the most likely to be useful and steps that could be taken to resolve the data problems permanently.

II. METHODOLOGY

This chapter addresses the two unique methodological problems we encountered in this study. First, it was necessary to ask: What are the relevant metrics for gauging the accuracy of independent estimates of procurement cost in the context of the acquisition process? Second, we need to examine the contention that use of the independent cost estimate as a basis for funding decisions is a “self-fulfilling prophecy;” that is, that funding a program to the independent cost estimate, while it may result in less cost growth, will result in a higher procurement cost.

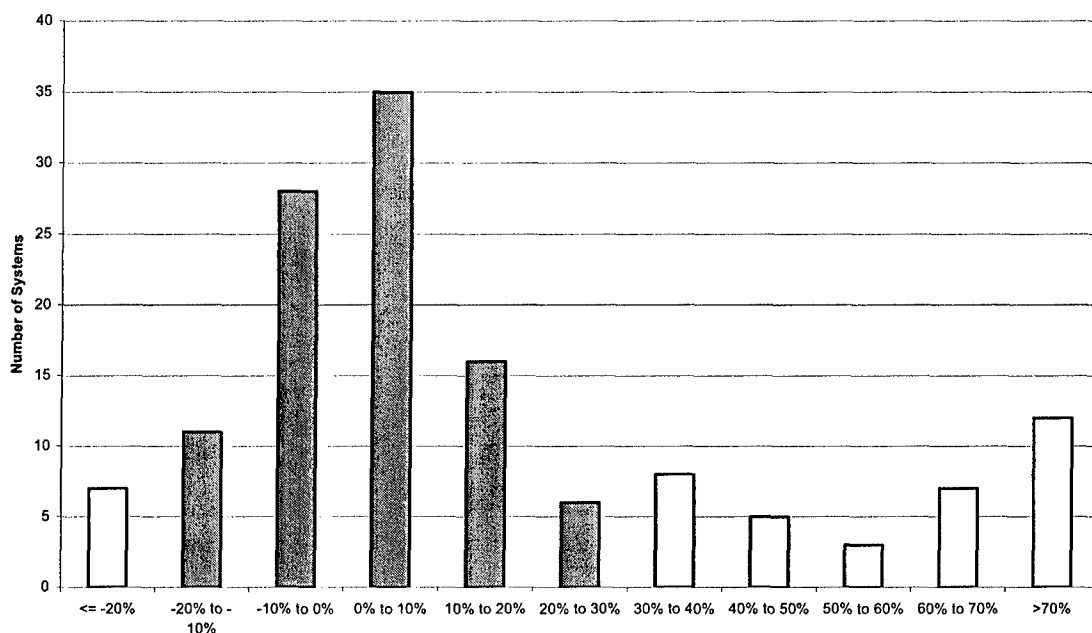
A. DISTRIBUTION OF PROCUREMENT COST GROWTH

It is useful to begin with a brief overview of the pattern of growth in procurement cost against the baseline approved when the program was authorized to begin EMD. (During the years from which our sample is drawn, approval to begin EMD for an MDAP occurred at MS II.) The relevant data are displayed in Figure 1. The data in this figure are taken from a database maintained by the OSD Office of Program Analysis and Evaluation (PA&E) and extracted from the Selected Acquisition Reports (SARs).⁹ Appendix A briefly describes how PA&E adjusted the SAR data for inflation, quantity change, and unforced changes in program content.

We use these data as the measure of actual growth in procurement cost from the MS II baseline.¹⁰

⁹ The PA&E cost growth database has been posted to a Web site. Inquiries should be directed to Director, Economic and Manpower Analysis Division, Office of Program Analysis and Evaluation, Office of the Secretary of Defense.

¹⁰ The PA&E data are taken from what are called the December SARs. These SARs report actual expenditures for years through the year of submission and budgeted amounts for the years of the Future Years Defense Program (FYDP). The programs we examine are far enough along that the data reported in their SARs provide a reasonable indication of the actual cost of the program adopted at MS II.



Source: PA&E cost growth database.

Note: The current estimate was restated on the basis of the total inventory objective approved at Milestone II and adjusted for unforced changes in program content.

Figure 1. Distribution of Adjusted Procurement Cost Growth from the Milestone II Baseline for 138 Major Weapon System Procurements

The PA&E database includes 138 systems that passed MS II during FY 1970–1997 inclusive. The cost growth for each system, measured against the MS II cost baseline, is:

$$\text{Cost Growth} = \left(\frac{\text{Current Estimate}}{\text{MS II Estimate}} - 1 \right) \times 100\%$$

Thus, cost growth of 50 percent, for example, means that the current estimate of procurement cost is half again what was forecasted at MS II.

PA&E states both the MS II estimate and the current estimate in the same year's dollars. The total procurement cost reported in the SAR reflects the Defense Department's plan as of the date of the SAR for the total quantity of the system to be procured. This quantity commonly differs from the total quantity approved at MS II. Consequently, PA&E adjusts the current estimate reported in the SAR to be the current estimated cost of buying the total quantity approved at MS II. Finally, the PA&E data attempt to adjust the estimates for changes in program content other than those forced by unrealistically optimistic MS II cost or schedule estimates.

For reasons stated below, we take the central portion of the distribution to be that between -25 percent and +25 percent cost growth. Ninety-one of the 138 systems lie in this interval.

Seven of the 138 systems underran their MS II procurement cost baselines by at least 25 percent. This feature of the data is not an artifact of PA&E's adjustments of the procurement cost growth data for content change, since 6 of the systems also showed negative growth of at least 25 percent before the adjustment.¹¹ It is also interesting to note that another 35 systems underran their MS II baselines by lesser magnitudes.

The right-hand side of the distribution is more heavily populated. Forty of the 138 systems experienced at least a 25 percent growth in procurement costs. These systems have an average procurement cost growth of 70 percent and account for over three-quarters of the total cost growth for the 138 systems.

B. METRICS

We now turn directly to the question of what metric or metrics should be used in gauging the accuracy of independent estimates of procurement cost. As a general proposition, the correct measure of accuracy of an estimate depends on what errors matter most to the decision maker who uses it. The context assumed here is a Defense Acquisition Board (DAB) MS II review of an MDAP, and the decision maker ordinarily will be the USD(AT&L).

To our knowledge, no USD(AT&L) has ever publicly identified the sort of measure he would want to employ in characterizing the accuracy of Service Cost Positions (SCPs) and Independent Cost Estimates (ICEs). It is possible, however, to infer plausible metrics from the context in which the SCPs and ICEs are used.

During the years from which our data were taken, the largely unwritten DAB "rules of the road" with respect to cost issues were as follows:

1. The SCP was the default funding option.
2. Cost issues were raised to the DAB on a "by exception" basis; that is, cost was not an issue unless someone with standing to do so raised it.
3. As a rule of thumb, cost was raised as an issue if the ICE for some major part of life-cycle cost (in particular, development, procurement, and operations and maintenance) was more than 10 percent above the SCP for that part of cost.

¹¹ See McNicol, "Cost Growth in Major Weapon Procurement Programs," Figure 4, p. 18.

4. The issue would be presented in terms of the specific elements of cost that accounted for the bulk of the difference between the ICE and the SCP (e.g., learning curve slopes for direct manufacturing labor, provision for initial spares, and business base assumptions).
5. If USD(AT&L) accepted the CAIG's arguments, an adjustment was made in the element of cost in question (e.g., the Service sponsoring the MDAP might be directed to fund to the CAIG's estimate of funding required for initial spares).

We noted in Chapter I that independent costing was introduced into the Defense Department primarily as a way of flagging instances in which the SCP for a major acquisition program was unreasonably optimistic. Viewed against this purpose, the DAB rules of the road on costing present the USD(AT&L) with two main types of cases. The first of these are the cases in which the ICE for procurement is roughly equal to the SCP. In this case, the ICE effectively endorses the SCP. Procurement cost, therefore, ordinarily would not be raised as an issue at the DAB, and the program would be funded on the basis of the SCP. The implicit question in this case is: How reliable a guide to the reasonableness of the SCP is agreement of the ICE? How frequently does the ICE fail to flag an SCP that proves, in retrospect, to have been unrealistically optimistic?¹²

The second main case is that in which the ICE for procurement is substantially higher than the SCP, and procurement cost is raised as an issue at the DAB. The principal question about the ICEs for this category of cases is: How many prove in retrospect to have been "false alarms," that is, cases in which the SCP turns out to have been a realistic estimate, although at the MS II DAB it was substantially less than the ICE?¹³

In summary, it is plausible to assume that a USD(AT&L) is mainly interested in two aspects of the accuracy of the ICEs:

1. How frequently the ICE fails to identify an SCP at MS II that proves to be extremely optimistic.

¹² As we will see, the ICE failed to detect an SCP that was unrealistically high in a few cases. These cases are not well known, however, and do not seem to have had any influence on expectations about the role of the ICE. The usual expectation is that the SCPs tend to be optimistic, and that the main job of the ICE is to flag unreasonably optimistic SCPs.

¹³ A third possibility is that the ICE for procurement is less than the SCP. This does occur; the MS II ICE for procurement was less than the SCP for four programs in our sample. In each of these cases, the difference between the SCP and the ICE is small, and in none of them was cost raised as an issue at the DAB.

2. How frequently the ICE incorrectly identifies the SCP as unrealistically optimistic.

Further progress requires providing some definition of the key terms used in these metrics. A good place to begin is with the notion of a “realistic” cost estimate. The term “realistic cost estimate” is not defined in statute or DoD regulations, but it usually is understood to mean a funding level that is adequate for the program to be executed but lean enough so that the program manager and the contractor can stay within the budget only by working hard at containing cost.

Note that this definition does not imply that a program that enters EMD funded on the basis of a realistic cost estimate will not experience any cost growth. First, cost growth post-MS II can occur for reasons unconnected to the realism of the MS II cost estimate, and all cost estimates have some inherent uncertainty. Second, as long as the program experiences only moderate cost growth, the MS II estimate will still be judged by most to have been realistic. The questions, then, are: What constitutes “moderate” cost growth? At what point is cost growth judged to be “extreme” and, therefore, to suggest that the initial estimate was unreasonably optimistic?

Statute provides help in identifying what amount of cost growth is acceptable to decision makers. In 1982 the Congress established requirements for reporting the cost growth experienced on individual MDAPs.¹⁴ The most relevant of these to the purpose of this discussion concerns average unit procurement cost (AUPC), which is defined simply as total procurement cost divided by total quantity procured.¹⁵ The Congress must be notified whenever the AUPC increase exceeds 15 percent of its level in the Acquisition Program Baseline. If AUPC growth exceeds 25 percent, USD(AT&L) must, in addition to notifying the Congress, certify that the acquisition program is essential to national security and that the revised estimate of the AUPC is reasonable. The statute further precludes obligating funds for any of the program’s major contracts if the required certifications are not provided to the Congress. Thus, cost growth of 25 percent can

¹⁴ Pub. L. 97-252 (Defense Authorization Act for Fiscal Year 1983), Title XI, Sec. 1107(a) (1), currently codified as 10 USC 2433. The amendment that led to this provision was introduced by Senator Sam Nunn and Congressman Dave McCurdy. The events that trigger these reporting requirements are often referred to as Nunn-McCurdy breaches.

¹⁵ AUPC does not include development costs or military construction costs. These are, however, included in the numerator of program acquisition unit cost (PAUC).

reasonably be taken as an indication of how much procurement cost growth is “too much” from the congressional point of view.

On this basis, we define growth in procurement cost of +25 percent or more, or -25 percent or less, as extreme cost growth. The lower limit of this range is arbitrary, as there is no DoD or congressional standard comparable to the + 25 percent for cost increases. Cost growth that falls inside these limits is judged to be acceptable. The ICE for each system in the sample is assigned one of three ratings:

- Too Low—the ICE is at least 25 percent below actual procurement cost;
- Acceptable—the ICE is within ± 25 percent of actual procurement cost;
- Too High—the ICE is at least 25 percent above actual procurement cost.

“Actual procurement cost” here means the procurement costs from the PA&E database (that is, procurement cost in constant dollars, restated on the basis of the total inventory objective approved at MS II, and embodying an attempt to adjust for content changes unrelated to the MS II cost estimates).

Statute provides the rationale for defining the Acceptable range of procurement cost growth to be ± 25 percent of the MS II estimate, but it is also sensible in terms of the uncertainty inherent in MS II cost estimates. The procurement estimates whose accuracy is reviewed in this study were made at the time that the program was just entering the EMD phase. At this point, many of the large features of the program that determine procurement cost had been selected. Development had not been completed, however; procurement expenditures were still several years in the future; and, once begun, procurement would continue for upwards of 10 years. Consequently, it is reasonable to expect a considerable degree of uncertainty in the estimates made at MS II.

In applying these ratings some information is necessary on the causes of the cost growth in the SARs, which record cost growth from all causes. Procurement cost growth that is not directly or indirectly related to inaccuracies in the MS II cost estimates clearly should be excluded when assigning the ratings to the ICEs. The PA&E data do this only partially. Note also that in assigning ratings to the ICE and the SCP, it is necessary to know what cost estimate was embodied in the MS II funding baseline.

We applied the ratings separately to ICEs for two subsamples distinguished on the basis of the DAB’s rules of the road on cost issues stated previously.

In the first subsample, the MS II ICEs for procurement were within 10 percent of the MS II SCPs.¹⁶ We can be reasonably sure in these cases that procurement cost was not raised as an issue with the DAB, and that the procurement portion of the MS II baseline was based on the SCP (which obviates for these cases any question about the ICE being a self-fulfilling prophecy). The DAB, in effect, treated the ICE as agreeing with the SCP for this subsample. The implicit premise was that agreement between the ICE and the SCP provided a solid basis for concluding that the program was realistically costed and would experience only modest cost growth. Ratings of the ICEs in this subsample, assigned as described above, answer the question: How often did the ICE fail to identify an unrealistic SCP that gave rise to extreme cost growth?

In the second subsample, the MS II ICEs for procurement were more than 10 percent above the MS II SCPs. It is very likely that procurement cost was raised as an issue at the DAB in these cases. The characteristic question for this subsample of cases is: Was the ICE a false alarm? That is, did the ICE incorrectly identify as unreasonably optimistic an SCP that in fact proved to be realistic?

It is reasonable to assume that all of the programs in the first subsample were funded on the basis of the SCP. This is not necessarily the case for programs in the second subsample, however. On rare occasions a program was funded entirely on the basis of the ICE, and the ICE more frequently influenced funding. These instances present the possibility, to which the discussion now turns, that partial or complete acceptance of the ICE was a self-fulfilling prophecy.

C. SELF-FULFILLING PROPHECIES

A proposal to base a program's budget on the ICE typically evokes considerable controversy, which not uncommonly includes the proposition that funding to the ICE is a "self-fulfilling prophecy." This phrase evokes an argument that starts with the fact that a weapon procurement program cost estimate is not simply a forecast, but is also presumably the basis on which the program's budget will be established. The program's budget is one of the tools a Government program manager (PM) uses to manage the

¹⁶ From the establishment of the CAIG through 1992, the CAIG's main role was to review the SCP. Consequently, the ICE differed from the SCP only to the extent that the CAIG objected to some part of the SCP and replaced it with its own estimate. Since 1992, the initiative has tended to shift to the CAIG, and some of its most constructive work is reflected in changes to SCPs made at an early stage, before they reach the DAB or the prior review of the readiness of the program for DAB review.

program. The assertion, then, is that the ICE should not be funded because it is too high to provide the PM adequate control over the contractor and those in the sponsoring Service who press for additions to the program. If so, by relaxing the budget constraint, funding the program to the ICE instead of the SCP might result in greater *cost*, although possibly in less *cost growth*.

As noted previously, procurement cost issues brought to the DAB ordinarily would concern a small number of subelements of procurement cost that accounted for the bulk of the difference between the ICE and the SCP. During some of the years represented in our sample, decisions to fund a program to some specific elements of the ICE were fairly common. It seems unlikely that such decisions would raise serious concerns about whether funding to the blended cost estimate was likely to be a self-fulfilling prophecy. This is so because the ICE was used only for specific elements of cost in these cases, and only then if the CAIG met the burden of proof the MDA required to conclude that some part of the SCP was unrealistic.

It clearly is necessary to confront the "self-fulfilling prophecy" hypothesis in those cases in which the MDAP is funded entirely on the basis of the ICE.

The principal counter to the argument that funding to the ICE is a self-fulfilling prophecy is that the ICE is a realistic estimate in the sense of the definition offered above: high enough for the program to be executable, but lean enough so that the program manager and the contractor can stay within the budget only by working hard at containing cost. Thus, proponents of the ICE would not agree that funding to it would cause procurement costs of the program to be higher than they would be if the program were funded to the SCP at MS II.

It is possible to establish in retrospect which side of such a disagreement was correct. Suppose, first, that the program was funded on the basis of the ICE. In the view of those who argued that funding to the ICE was a self-fulfilling prophecy, that means the program was overfunded. There are two clear markers of an overfunded program. First, if a program is executed successfully after funding was removed during the budgeting process without removing program content, we could definitively conclude that the program had been overfunded. (The program would, after removal of the extra funding, show negative cost growth against its MS II baseline.) Second, addition of program content without any increase in funding would be an equally conclusive marker of program overfunding. (This is what usually is envisioned as the way in which a self-fulfilling prophecy is fulfilled.) There is a good possibility that each of these markers

(when they occur) can be recognized in available data. If either of these situations occurs, the program was clearly overfunded, and if the slack were taken up by addition of content, we can conclude that funding to the ICE was a self-fulfilling prophecy.¹⁷ In contrast, if the ICE in question were a realistic estimate (rather than too high), we would expect a program funded to the ICE to experience comparatively little increase in the cost of providing the content approved at MS II.

The other, more commonly encountered, possibility is that the program is funded on the basis of the MS II SCP. If the SCP is a realistic estimate, we would again expect to find that any substantial cost growth was clearly traceable to some cause other than the MS II cost estimate. If that is observed, we could conclude that funding to the ICE would have been an error and a self-fulfilling prophecy unless the excess funds were removed at some point after the MS II review. If, however, the SCP was in fact unrealistically optimistic, we would find that funds were added, while program content remained unchanged, or content removed so that the program could be executed within MS II funding levels.

These observations are summarized in Table 2. We assumed that either the SCP or the ICE is funded in any particular case. We do not observe what would have happened had the other choice been made. Consequently, to evaluate the self-fulfilling prophecy argument, we ideally would have a set of programs for which the ICE was substantially above the SCP at MS II, some funded on the basis of the SCP and others funded on the basis of the ICE. Unfortunately, as discussed below, our sample includes only one program funded entirely on the basis of the ICE at MS II.

Table 2. Principal Markers of Realistically and Unrealistically Funded Procurement Programs

Program	Realistic	Unrealistic
ICE-Funded	Little growth in cost of MS II content	Funds removed; or content added, funding unchanged
SCP-Funded	Little growth in cost of MS II content	Funds added, content unchanged; or content removed, funding unchanged

¹⁷ A third possible marker would be increased organizational slack. This would be hard to document with readily available data, and, given the Services' attention to program management, probably is not often a major factor.

D. CONCLUDING COMMENTS

The conclusions of this chapter are the identification of the markers of a self-fulfilling prophecy (Table 2 previously) and the two metrics against which the accuracy of the ICEs and the SCPs will be assessed:

- How frequently the ICE fails to identify an SCP at MS II that proves to be extremely optimistic.
- How frequently the ICE incorrectly identifies the SCP as unrealistically optimistic.

The following chapter describes the data used in the assessment. We then pick up the thread of the discussion above in Chapter IV, which presents our results.

III. DATA

We included in our initial sample all MDAPs that passed MS II between 1987 and 1997 inclusive. In a first sweep of the data, we inadvertently picked up 2 MDAPs that passed MS II in 1998, and 6 that passed MS II in 1985. After looking at these cases, we decided to retain all of them in our initial sample, which then included a total of 63 MDAPs.

The choice of the interval 1985–1998 reflects a balance of two considerations. First, the earlier the year considered, the sparser the data available and the more difficult they are to collect. Second, the further past MS II a system is, the more certain are its actual costs. By cutting the sample off at 1998, we insure that all systems in the sample are at least 3 years into EMD (because the most recent cost data we had when this study began was for 2001).¹⁸

The remainder of this chapter briefly describes each of the main sets of data the study required, indicating our sources, the main shortcomings of the information we used, and what we did in response to those shortcomings.

A. SCP AND ICE

We attempted to locate for each of the 63 MDAPs in the sample the CAIG estimate made to support the MS II decision and the corresponding SCP. PA&E's Resource Analysis Deputate has retained copies of many of the memoranda that the CAIG has through the years provided to successive MDAs in support of milestone reviews of major systems.¹⁹ The CAIG permitted us access to such MS II CAIG reports as they had retained for the MDAPs in the initial sample. We collected both development and procurement cost estimates, but because of time and resource constraints our analysis was limited to the procurement cost estimates.

All of the CAIG reports from the start of FY 1994 on provide numerical values for the CAIG and the SCP procurement estimates. Those before FY 1994 do not always do

¹⁸ We used a cost growth database maintained by PA&E. PA&E did not update the data in FY 2002, and the update for FY 2003 was released after this study was substantially complete.

¹⁹ The PA&E Deputy Director for Resource Analysis serves as the chairman of the CAIG.

so. In particular, some of the CAIG reports from earlier years provide only a qualitative assessment of the degree of risk in the SCP, without providing a specific CAIG estimate.

The change in the record reflects the assumption by the CAIG in February 1992 of sole responsibility for producing the independent cost estimates for MDAPs required by statute. (The CAIG was not fully staffed to perform this function until early FY 1994.) As noted previously, from the formation of the CAIG in FY 1972 through mid-FY 1992, the CAIG shared responsibility for producing these estimates with cost groups located in the secretariats of the military departments. During these years, the CAIG operated by reviewing and selectively replacing parts of cost estimates prepared by the military departments' headquarters-level parametric cost groups. In some cases, the CAIG provided the DAB chairman only a qualitative assessment of the SCP.

We found in the CAIG reports the ICEs and SCPs for 30 of the 63 MDAPs in our initial sample. Five of these 30 MDAPs proved to be unusable, for a different reason in each case. The remaining 25 MDAPs constitute our usable sample. We could not locate a CAIG report for 15 MDAPs and we therefore do not have either an ICE or an SCP for them. We had either the ICE or the SCP but not both for 18 MDAPs.

Figure 2 is a histogram of the difference between the ICE and the SCP for the 25 MDAPs for which both were available.

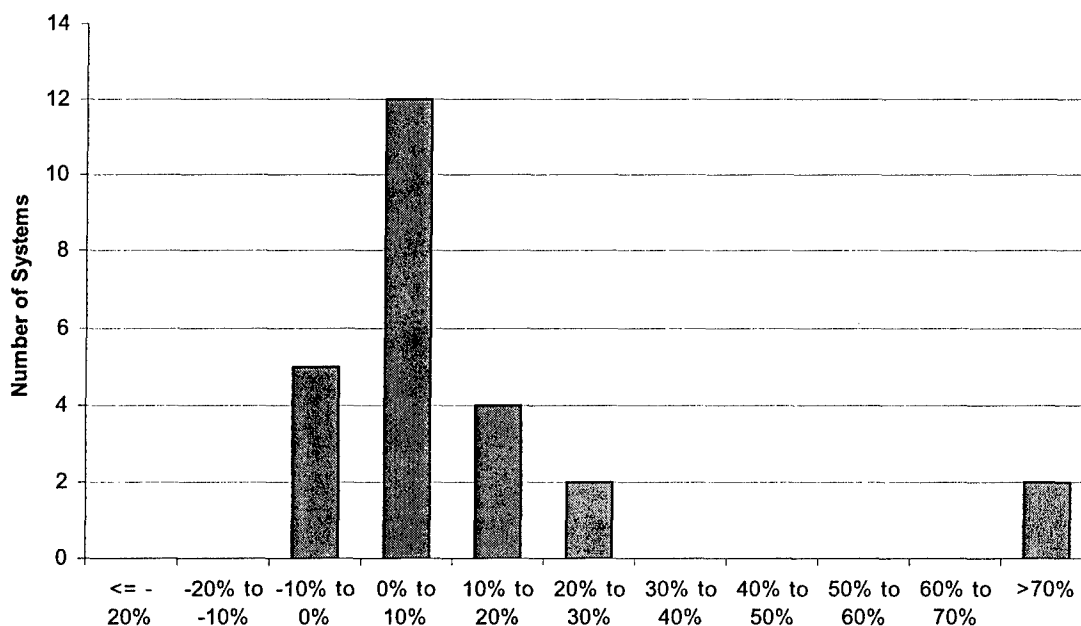


Figure 2. Histogram of the Percentage Difference between the Independent Cost Estimate for Procurement and the Service Cost Position for Procurement for 25 Major Defense Acquisition Programs

B. WHAT WAS THE BASIS OF THE FUNDING DECISION?

The next question we need to answer is: What was used as the basis of the funding decision at MS II? This question proved to be very difficult, and the answers we got are not all equally defensible.

We used five different approaches to establishing the basis of the MS II funding decision:

1. *Acquisition Decision Memorandum (ADM)*. We have the ADM for all but 4 of the systems in our usable sample. In no case, however, did the ADM indicate which of the estimates was used as a basis for the MS II procurement funding decision.
2. *Defense Acquisition Board (DAB) "Rules of the Road."* The 25 MDAPs in our usable sample contain 17 instances in which the ICE for procurement cost differed from the SCP by no more than 10 percent. As noted in the preceding chapter, cost ordinarily would not be raised as an issue in these cases, and the SCP was the default funding option.²⁰ Accordingly, in these 17 cases we believe the SCP was the basis for funding the program.
3. *Acquisition Program Baseline (APB)*. The APB adopted at MS II records the MDA's decision on funding. Hence, comparison of the SCP and the ICE to funding streams in the APB should provide a reasonably clear indication of which estimate was funded. We used the APB to get a reasonable indication of which estimate was funded at MS II for four of the MDAPs in our usable sample.²¹
4. *Selected Acquisition Report (SAR)*. The first SAR filed after MS II is passed ordinarily provides procurement funding data. These data might, then, be used in place of the APB to establish which estimate was funded. We used this approach for two of the MDAPs in our usable sample.
5. *Recollections of the CAIG staff*. In two cases we based our assessment on the recollections of CAIG staff who had been closely involved because none of the other methods produced a usable answer.

²⁰ The decision on whether to raise cost as an issue was ordinarily made by the chairman of the Overarching Integrated Product Team. The chair of the OSD CAIG could, however, independently bring the issue to the USD(AT&L).

²¹ In many of the other cases, the APB funding did not match either the ICE or the SCP. The likely explanation of these mismatches is that the program content approved at MS II was substantially different from that assumed in the SCP and the ICE. In such a situation, the SCP (or occasionally the ICE) would be used to compute the cost implications of the program alterations.

C. PROCUREMENT COST GROWTH DEFINED

The SARs record procurement cost growth for the program in both then-year dollars and program base-year dollars. The SCP and the ICE also should be reported in then-year and program base-year dollars. Hence, computation of an overall figure for procurement cost growth is just a matter of division. Obtaining an “apples-to-apples” comparison of cost revealed in procurement experience and the MS II baseline is ordinarily much harder. To get such a comparison it is necessary to take account of:

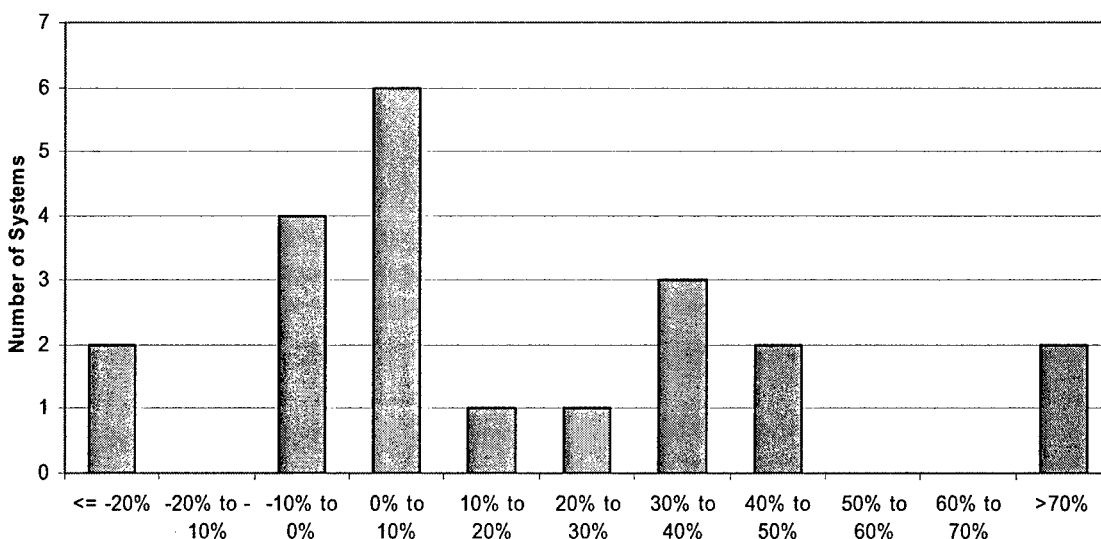
1. Changes in the total quantity purchased.
2. Changes—addition or removal—of content from the MS II program.

Making these corrections is seldom a matter of posting clearly visible cost changes to well-defined categories. Rather, a considerable amount of fact finding and analysis is typically required to unravel the implications for cost of the changes that almost invariably take place in a program after it passes MS II.

The analytical sophistication that SARs bring to resolving these issues is uneven, and the SARs’ treatment of them is not necessarily entirely consistent from one MDAP to another or over time for a given MDAP. Recognizing that, in the mid-1990s PA&E undertook a long-term effort to recast systematically the cost variances reported in the SARs. To date, that has been accomplished for 138 systems.

We took our procurement cost growth figures from the database PA&E makes available to DoD users. In these data, variances reported in the SARs are sorted between those due to unforced changes in program content and those that represent an increase in the cost of the content approved at MS II. (As noted previously, the adjustments PA&E makes to the SAR data are briefly described in Appendix A.) We used the PA&E data rather than attempting our own sort of the SAR variances. As noted below, the PA&E data, although a large step forward, have important shortcomings in some cases.

Figure 3 is a histogram of the PA&E cost growth data for 21 of the systems in our usable sample. Usable cost growth data for the other four systems is not yet available. The following chapter notes how the ICEs were assessed in these four cases.



Source: PA&E cost growth database.

Notes: In constant program base-year dollars. The current estimate was restated on the basis of the total inventory objective approved at Milestone II and adjusted for unforced changes in program content.

The PA&E database does not yet contain procurement cost growth for two of the systems in the sample, and for two others, the relevant figures for the purposes of this study are not those in the PA&E database. See Section B of Chapter IV for the basis of rating the ICEs for these four systems.

Figure 3. Distribution of Adjusted Procurement Cost Growth from the Milestone II Baseline for 21 Major Defense Acquisition Programs

D. CAUSES OF COST GROWTH

DoD weapon system cost estimators are responsible for making realistic estimates of the costs of programs that have been specified by others. That is to say, the estimates prepared by cost estimators are crucially dependent on technical and programmatic assumptions over which they have little or no say. There are some gray areas; cost estimators should recognize—and provide corrections for in their estimates—some types of unrealistic program assumptions and some likely execution problems. But, without trying to fix the boundaries of these exceptions, it is clear that they are exceptions—cost estimators generally are not equipped to do engineering analyses of proposed programs or to assess the capabilities of potential contractors. Moreover, cost estimators are simply advisers to the MDA, who makes the decision on what estimate to use in funding decisions.

It is therefore not safe to assume that high procurement cost growth implies poor cost estimation at MS II. There are several other possible explanations:²²

1. A realistic cost estimate was available to the MDA but was not used.
2. The program approved at MS II rested on an unrealistic technical or programmatic assumption, which the cost estimators could not reasonably be expected to identify.
3. The acquisition was pursued with a badly flawed contracting strategy.
4. The program was poorly executed by the contractor or by the Defense Department (defined to include instances of exceptional budget instability).
5. The program was changed in a major way after it had been in EMD for some time, for a reason unrelated to the MS II cost estimate (for example, because of a change in the threat).

The adjustments made by PA&E ideally would remove the last of these, but a recent IDA study indicates that in some cases the adjustments made may be well short of what is required.²³

We could not in the context of this project do the thorough review of the cost estimates and history of the MDAPs in the sample necessary to get a solid assessment of the explanation of the instances of large positive or negative cost growth in the programs included in the sample. Such assessments, however, obviously are a necessary step in an assessment of the accuracy of the MS II ICEs. The tentative assessments we report rest on the recent IDA study of cost growth cited above, on program narratives in the SARs, and on the experience of those at IDA involved in this project. It is highly likely that more detailed study would at a minimum refine and, perhaps, reverse some of our assessments of what post-MS II cost growth is chargeable to the MS II cost estimates.

E. CONCLUDING COMMENTS

The four sets of data used in this study differ greatly in the nature and degree of uncertainty in them. For each MDAP we either have the record of the ICE and SCP for procurement or we do not. There is no uncertainty in this. The basis for the MS II

²² Substantially this list of mechanisms also must be evoked to explain underruns. Note in addition that a deliberate overestimate of a system's likely cost provides a way of holding a reserve in the acquisition account. Somewhat more visible ways of holding such reserves usually risk having the funds reallocated either during the DoD resource allocation process or during congressional deliberations on the President's budget submission for the Defense Department.

²³ McNicol, "Cost Growth in Major Weapon Procurement Programs," pp. 35-40.

funding decision is also in principle knowable with certainty; the source of uncertainty in this case is simply shortcomings in the records we could locate and review within the confines of this study. Computing growth in the cost of the program approved at MS II is a formidable undertaking; it is effectively a matter of "backing out" the costs of unforced changes made in the program after MS II. This requires a great deal of information and presents some substantial estimating challenges. Nonetheless, such tasks are frequently done and are well understood.

Against this background, the assignment of the reasons for cost growth stands out as much more challenging. This challenge cannot be avoided in an assessment of the accuracy of independent weapon system cost estimates in that use of an inaccurate cost estimate at MS II is but one of several ways in which major procurement cost growth can occur. Assignment of procurement cost growth to these causes plays a large role in the conclusions on the accuracy of the independent estimates of procurement costs. It is therefore important to repeat that within the resource limitations of this study we could not undertake the detailed examinations of the individual cases that would be necessary to have reasonably high confidence in assignment of the causes of cost growth.

IV. RESULTS

The first two sections of this chapter present our assessment of the MS II procurement ICEs and SCPs for the 25 usable MDAPs in our sample against the two metrics identified in Chapter II. The third section examines the one case in the sample for which procurement at MS II was entirely funded on the basis of the ICE for evidence of a “self-fulfilling prophecy” effect.

A. FREQUENCY OF FAILURES TO IDENTIFY AN UNREALISTICALLY OPTIMISTIC SCP

Procurement cost probably was not raised as an issue at the MS II DABs for 17 of the 25 MDAPs in our usable sample.²⁴ Following the rules stated in Chapter II, the SCPs for these 17 programs were scored as Acceptable if the actual cost of the system—normalized for inflation, change in inventory objective, and program content—was within ± 25 percent of the MS II baseline. Cost growth was in this range for 10 of the 17 systems. We scored these as Acceptable without further investigation.

Cost growth was more than 25 percent for 6 of the 17 systems, and less than -25 percent for one system. In these 7 cases, we asked whether something other than adoption of an unrealistic estimate at MS II was at work. A recent IDA study suggests that other factors were major causes of procurement cost growth in 4 of these 7 cases.²⁵ (Each of these programs experienced large cost overruns.) The SCPs for these 4 systems therefore were scored as Acceptable (for a total of 14 in the Acceptable category). We found no explanation for the other 2 cases of cost growth of more than 25 percent, which were therefore scored as Too Low, or for the underrun of more than 25 percent, which was scored as Too High.

²⁴ These are the systems for which the procurement ICE was within 10 percent of the SCP for procurement. It is worth noting that for 4 of the 17 the ICE was less than the SCP, although never by more than 2 percentage points.

²⁵ See McNicol, “Cost Growth in Major Weapon Procurement Programs,” especially Chapter 4 and Table 14, p. 69.

In the 17 cases under consideration, the DAB effectively took the rough agreement of the ICE with the SCP to indicate that the SCP was realistic. For that reason, for these MDAPs, we assigned the ICE to the same category as the SCP.

Our results for this set of cases are presented in Table 3. In 80 percent (14 of 17) of the cases in which the ICE was roughly equal to the SCP, the SCP proved to be a realistic estimate by the criteria of this study. In two cases the ICE failed to identify an SCP that proved to be unrealistically optimistic, and in one case the ICE failed to identify an SCP that proved to be unrealistically high.

Table 3. Assessed Accuracy of the ICEs for 17 MDAPs for Which Procurement Cost Was Not Raised as an Issue at the DAB Milestone II Review

	Too Low	Acceptable	Too High	Total
SCP	2	14	1	17
ICE	2	14	1	17

B. FREQUENCY OF FALSE ALARMS

The ICE exceeded the SCP by more than 10 percent for 8 of the 25 systems in the sample. Cost, or aspects of the program closely tied to cost, probably were an issue at the MS II DAB in each of these cases.

To score these 8 cases, it is necessary to identify the cost estimate on which the MS II program funding baseline was established.²⁶ To see why that is so, suppose that (1) observed cost growth relative to the MS II baseline is 10 percent and (2) the ICE is 35 percent above the SCP. If the program was funded on the basis of the ICE, we would score the ICE and the SCP as follows:

- SCP: Too Low—the SCP is 45 percent below actual cost;
- ICE: Acceptable—the ICE is within 10 percent of the actual cost.

Note that the cost growth attributed to the SCP (10 percent actual growth plus the 35 percent the SCP was below the ICE) is what we assume cost growth would have been

²⁶ It is also necessary to score the accuracy of the ICEs and SCPs in the case, considered in the preceding section, in which procurement cost is not raised as an issue at the DAB. In that case, however, it is safe to assume for the years from which our sample is drawn that the program was funded on the basis of the SCP.

had the program been funded to the SCP. Suppose, instead, that the MS II baseline was the SCP. Then our scoring would be:

- SCP: Acceptable—the SCP is within 10 percent of the actual cost;
- ICE: Too High—the ICE is -25 percent below actual cost.

The point of this example is simply that misidentification of the basis of the MS II baseline probably will result in major errors in scoring the accuracy of the SCP and the ICE.

Five of the 8 programs were funded entirely on the basis of the SCP. One of the programs was restructured in large part because of the ICE, and the restructured program subsequently funded on the basis of the SCP. Of the remaining 2 programs, one was partially and the other entirely funded on the basis of the ICE. Consideration of whether the one case funded entirely on the basis of the ICE proved to be a self-fulfilling prophecy is postponed to the following section.

Four of the 8 cases in this subsample can be scored by a straightforward application of the quantitative rules of Chapter II.²⁷ In 2 of the cases, the ICE was Too High. In one of these, actual cost growth placed the SCP in the Acceptable range; in the other, it was, like the ICE, in the Too High category, but more nearly accurate. In 2 cases the ICE proved to be Too Low; that is, relative to the ICE, actual cost growth exceeded 25 percent. These are somewhat remarkable cases in that the ICE was well above the SCP for both of them.

The 4 remaining cases could not be scored strictly by the quantitative rules of Chapter II. In narrow terms, the reason in each case was that the relevant cost growth figure cannot be computed with available data. In each of these cases, we scored the ICE as Acceptable and the SCP as Too Low, because the ICE presented a major reason why the SCP would prove to be flawed, and subsequent evidence indicates the ICE prediction was substantially accurate. In these cases, "Acceptable" does not mean that the MS II ICE or SCP was within ± 25 percent of the actual cost. Rather, it means that the ICE correctly pointed out that the SCP was unrealistically optimistic in some major respect.

Table 4 summarizes our assessments for the 8 cases in which cost was most likely raised as an issue at the MS II DAB. The ICE was a false alarm in 25 percent of these

²⁷ In one case, we classified the ICE as Too High, even though by the rules of Chapter II it fell just within the Acceptable range. We adopted this classification because it seemed warranted by the facts of the case.

cases (2 of 8). The ICE correctly identified an SCP that was unrealistically low in the other 75 percent of the cases (6 of 8). The weight placed on this assessment, however, must be tempered by the fact that 2 of the 4 programs for which the ICE was scored as Acceptable used unprecedented contracting strategies, another one was cancelled while in EMD, and the last was restructured after an initial unsuccessful attempt to gain MS II approval.

Table 4. Assessed Accuracy of the ICEs for 8 MDAPs for Which Procurement Cost Was Raised as an Issue at the DAB Milestone II Review

	Too Low	Acceptable	Too High	Total
SCP	6	1	1	8
ICE	2	4	2	8

C. SELF-FULFILLING PROPHECY?

As previously noted, the ICE was unequivocally used as the basis for funding for only one of the 25 MDAPs in our sample. In this case, the ICE procurement estimate was nearly double that of the SCP. The question for this case is whether funding the ICE was a self-fulfilling prophecy.

Significant content was added after MS II to the MDAP in question, but not within established funding. Rather, the added content came with increased funding. Moreover, after taking account of the content additions and quantity reductions, the ICE procurement estimate proved to be 20 to 25 percent too low, even though it was nearly double that of the SCP at MS II. This case, then, does not provide an example of a program that was overfunded.

We argue in Chapter II that funding a program to a cost estimate that is a blend of the SCP and the ICE usually will not raise a concern about whether doing so is a self-fulfilling prophecy. Unfortunately, our sample does not provide any test of this conclusion, since the one program apparently funded to a blended estimate was cancelled before it went into production.

D. CONCLUDING COMMENTS

Our results on the accuracy of the 25 ICEs we had sufficient data to score can be summarized as follows:

- The ICE failed in about 20 percent of the relevant subset of cases to identify an SCP that proved to be unrealistically low or high.
- The ICE was a false alarm in about 25 percent of the relevant subset of cases.

These conclusions provide a rough guide on the reliability of ICEs for use by USD(AT&L) and other senior officials involved in the acquisition of major weapon systems. It is important to note, however, that our conclusion on the accuracy of the SCPs and ICEs indicate the DAB cannot rely on either alone, and USD(AT&L) must be prepared to resolve major cost issues.

V. RECOMMENDATIONS

This paper's main results on the accuracy of the 25 ICEs we had sufficient data to score can be summarized as follows:

- The sample includes 17 cases in which procurement cost was not raised as an issue at the DAB; the ICE failed to identify the 2 cases in this subsample that proved to be unrealistically low and the 1 case that proved to be unrealistically high.
- The ICE was a false alarm in 2 of the 8 cases in which cost was raised at the DAB.

The study is not broad enough to warrant conclusions on whether this accuracy rate is acceptable or what might be done to improve it.²⁸ Accordingly, the recommendations offered are prompted by problems encountered in doing the study rather than by the results.

This study offers nine recommendations directed towards providing further insight into the accuracy of ICEs and SCPs. These recommendations are grouped under the following three general topics: conducting further studies, establishing a process to track cost issues, and capturing key funding data.

A. CONDUCTING FURTHER STUDIES

Recommendation 1: Expand the Sample

Including more MDAPs in the sample would increase confidence in the conclusions reached. For many of our initial sample of 63 MDAPs, doing so would be inordinately difficult. For others, only modest additional effort would permit them to be included in the usable sample.

²⁸ These questions are addressed in McNicol, "Cost Growth in Major Weapon Procurement Programs." See, in particular, Chapters 1 and 7.

Recommendation 2: Extend the Study to Development Cost

This study examined the accuracy of only the independent estimate of procurement costs. The exclusion of the independent estimates for the costs of EMD is a major limitation, and it would be of considerable interest to extend the analysis to them.

Recommendation 3: Examine in Detail the Causes of Cost Growth in Selected Cases

The largest uncertainty in this study's quantitative findings lies in the attribution of large cost growth in several cases to causes other than an unreasonably optimistic MS II cost estimate. For 4 MDAPs in the sample, the following was true: the ICE and the SCP were within 10 percent of one another; the program later experienced more than 25 percent growth in procurement cost; and some evidence suggests that a major part of the cost growth was caused by a factor other than a faulty MS II cost estimate. It would be useful to look at each of these 4 cases in detail. If in each case the attribution is in error, the Acceptable percentage when procurement cost was not raised as an issue at the DAB falls to 60 percent (from 80 percent).

Recommendation 4: Examine MDAPs That Underran Their Milestone II Cost Baseline

Only 2 MDAPs in our usable sample underran their MS II cost baseline, and in both cases the underrun exceeded 25 percent. The larger PA&E cost growth database contains 19 systems that apparently underran their MS II cost baselines by at least 20 percent. If the program content approved at MS II was procured in these cases, the underruns are conclusive evidence that the program was overfunded at MS II; that is, that the funding at MS II was based on a cost estimate that was too high. Accordingly, these cases of underruns might provide a good test of the main premise of the self-fulfilling-prophecy hypothesis. Addition of program content—that is, increased quantity or improved performance—within the MS II cost baseline would also be conclusive evidence of overfunding. Content additions would be harder to identify than cost underruns, but the SARs should permit identification of quantity additions and improved performance against established requirements.

Recommendation 5: Examine the Effects of Strengthening the CAIG

It also might be of interest to examine the effects of the substantial changes in the independent costing process made around 1994 on accuracy of ICEs. It was then that MDAPs coming to the DAB for a milestone review were required to provide the CAIG a

description of the proposed program in enough detail to permit it to be estimated.²⁹ At roughly the same time, the CAIG assumed sole responsibility for preparation of the ICEs required by statute, CAIG access to program office and contractor information was considerably increased, and the time allowed for the preparation of the ICE significantly increased.³⁰

B. ESTABLISHING A PROCESS TO TRACK COST ISSUES

Recommendation 6: Maintain Validated Cost, Schedule, and Performance Tracks

This paper is part of a larger, ongoing project intended to capture acquisitions lessons learned on a rolling basis. The larger project includes development of an assessment of how well each MDAP did against the cost, schedule, and performance baselines established when the program enters the System Development and Demonstration (SDD) phase. These data are to be included in an archive of basic program documents from the SDD and the Full-Rate Production DABs. The difficulties this paper encountered in developing cost tracks against the MS II baseline provide evidence of the need for such a collection of information.

Recommendation 7: Track Cost Issues in Ongoing MDAPS

Substantial cost issues came up in connection with the MS II DAB review for 8 of the 25 MDAPs in our usable sample. Few of these were narrow questions about costing approaches or data of interest primarily to those who do weapon system costing; most were questions about such matters as acquisition strategy, the outcomes that would be provided by certain contractual vehicles, and policy on funding at MS II. These cases provide potentially useful sources of acquisition lessons learned. They also would be comparatively accessible if analysis were done as the program proceeds, but usually would be much more time-consuming if picked up with even the lapse of a year or two.

²⁹ The CAIG began requesting a Cost Analysis Requirements Description (CARD) about 1990, but the requirement to do so was not written into the regulations for the DAB process until 1994.

³⁰ McNicol, "Cost Growth in Major Weapon Procurement Programs," pp. 46–49, provides evidence that the strengthening of the CAIG around this time, all else equal, resulted in lower growth in procurement cost. "All else" was not equal, however, and cost growth for the years 1993–1997 inclusive was higher than it had been over the preceding 20 years.

C. CAPTURING KEY FUNDING DATA

The presumption is that the SCP, the ICE, or some blend of both is used to establish the funding approved at MS II and captured in the APB. We were therefore somewhat surprised to find that the funding in the APB was significantly lower than either the SCP or the ICE in 14 of the 25 MDAPs in our sample.³¹ Moreover, the funding in the APB did not always match that in the SAR, and it is our understanding that the funding in the SAR is sometimes significantly different from what is in the DoD program submitted to the Congress.

These comments point to two distinct problems: (1) the lack of a record of the basis on which the MS II funding was determined and (2) the mismatch between funding in the APBs, the SARs, and the defense program.

Recommendation 8: Record the Basis of the Milestone B Funding Decision

Since the early 1990s, the presumption has been that the programming and budgeting processes will defer to costing decisions made by USD(AT&L) in the acquisition process. Clearly, however, compliance with this presumption is apt to be loose if it rests in many cases on the imperfect recollections of participants. It would, therefore, be desirable to establish a vehicle for recording the link between the cost estimates presented to the DAB and the funding captured in the APB. This might be done in a brief annex to the APB, prepared by the Director, Acquisition Resources and Analysis, in coordination with the PM and the CAIG chair.

Recommendation 9: Make Funding and Quantity Decisions Captured in the APB Mandatory Guidance in the Joint Planning Guidance

Mismatches between funding in the APBs, the SARs, and the defense program can easily arise because the decisions made in the acquisition process on funding and annual quantities to be procured are not infrequently revisited in the resource allocation process. Within the new Planning, Programming, Budgeting, and Execution System, the appropriate remedy to this problem would be to include the funding and quantity profiles of the APBs as mandatory guidance within the Joint Programming Guidance. Changes would then be effected annually by Program Change Proposal (PCP) or Budget Change

³¹ APB funding was considered significantly lower than the SCP if the difference was greater than 10 percent.

Proposal (BCP). Any PCPs or BCPs accepted would have to be reflected in the relevant APBs, of course.

D. CONCLUDING COMMENTS

Our nine recommendations are roughly ordered from least to most difficult. The standard of difficulty used is not one of how analytically challenging the recommendation would be, but rather how difficult it would be to implement.

Achieving Recommendations 1 through 5 (further studies) would depend on data availability. There is no doubt, however, that each could be undertaken, and there is a reasonable prospect that each could shed some further light on the accuracy of ICEs.

Recommendation 6 (maintain validated cost, schedule, and performance tracks) is difficult primarily because it would require a large effort. Recommendation 7 (track cost issues) would require less effort, but its purpose is as much quality assurance as analysis, and hence might raise potentially difficult issues of organizational authority and accountability.

Recommendation 8 (recording the basis of the Milestone B funding decisions) would not require a great deal of additional effort. Moreover, we do not see any reason why it would encounter strong opposition.

Recommendation 9 (make funding and quantity decisions in the Joint Planning Guidance) would not require any major changes in DoD regulations. Implementing it would essentially be a matter of directing that USD(AT&L) send the APBs to the Director of PA&E, and that the Director of PA&E ensure that they are included as mandatory guidance in the Joint Planning Guidance. In fact, this change would arguably simply codify current guidance. This simple change, however, would require approval (through a BCP or PCP) of changes that the Services now often make themselves in formulating the Budget Estimate Submissions and Program Objective Memoranda. It is, then, a recommendation with consequences that go considerably beyond efficient capture of funding data required to evaluate the accuracy of the ICEs and SCPs.

APPENDIX A: DEFLATION, QUANTITY NORMALIZATION, AND ADJUSTMENT FOR CONTENT CHANGES

The procurement cost growth data used in this study were downloaded from a Web site OSD PA&E maintains. Some documentation of the data is provided on the opening page of the site. This appendix (excerpted from an earlier IDA publication¹) provides additional information based on notes provided to the author by PA&E staff.²

DEFLATION

The SARs, starting in 1974, report costs (including cost variances) in both program base-year dollars and then-year dollars. The data posted on the PA&E site, and used in computing cost growth, are in constant dollars. That is, the required reduction of the data to base-year dollars was done by the program offices in preparing the SARs for 1974 and subsequent years. A recent RAND Corporation report provides a useful overview of the evolution of SAR procedures for deflation and their shortcomings.³

QUANTITY NORMALIZATION

The same RAND study lays out the three methods available for normalizing the SAR cost data for changes in the inventory objective, that is, in the total quantity of the system procured.⁴ The PA&E data are normalized using the third of these, which is the most demanding in terms of data requirements and also the most satisfactory.

The first step is the classification of the posted SAR cost variances (in program base-year dollars) as “quantity-related” or “not quantity-related.” The dollar magnitude of a posted variance reflects the inventory objective as of the time the SAR is filed. A

¹ David L. McNicol, “Cost Growth in Major Weapon Procurement Programs,” Institute for Defense Analyses, Paper P-3832 (Nonstandard), 2004, Appendix C and pp. 18–20.

² The data used were those posted on the site as of about March 10, 2003.

³ Paul G. Hough, “Pitfalls in Calculating Cost Growth from Selected Acquisition Reports,” Chapter 6, RAND Corporation, N-3136-AF, 1992.

⁴ Hough, “Pitfalls in Calculating Cost Growth from Selected Acquisition Reports,” Chapter 7.

variance is judged to be "quantity-related" if it would be larger for a larger inventory objective (and smaller for a smaller inventory objective). Many elements of development cost are not quantity-related, but most other cost variances are. The quantity-related variances are cumulated.

The second step is the estimation of a learning curve for each system in the sample. This is done by PA&E staff. In most cases the learning curve slope is estimated by regressing cumulative output on cumulative cost as reported in successive SARs for the system. (For example, if a system ended after being in production for 15 years, the SARs for the system provide 15 pairs of cumulative output-cumulative cost observations.) The regression was not used if the estimated learning curve parameter was unreasonable (in particular, greater than one) or the R^2 was less than 0.7. In those cases, a learning curve slope was assumed based on the most pertinent information available (e.g., the slope characteristic of other systems in the commodity class).

The final step was the mechanical one of adjusting the cumulated quantity-related variance for the change in the inventory objective. This step is actually accomplished in the database software PA&E uses.

ADJUSTMENT FOR CONTENT CHANGES

It is a long-standing criticism of the SAR variance categories that they do little to reveal the underlying causes of the cost growth that is recorded in them. For example, procurement cost may have increased because at MS II the Service made (and placed on contract) an unrealistic assumption about what physical attributes the system needed to have to satisfy the requirements that had been established for it. Alternatively, the increase might be explained by a decision to buy a more capable system or to upgrade a system that had been in the field for years. Similarly, a quantity variance can occur when some change in circumstances—the threat the system is to meet, for example—leads to a reduction in the quantity of the system that the Service decides to buy, or it may reflect a revisit of the original quantity decision in the light of a considerably higher cost, which in turn may reflect the adoption of an unrealistic cost estimate at MS II.

The PA&E data used in this study take a first, limited step towards correction of this shortcoming of the SARs. They do this by separating for each system the amount of cost growth due to "mistakes" and the amount due to "decisions." These labels are italicized in what follows to emphasize that they stand for criteria for categorizing the data. In general terms, the *decisions* bin is intended to capture the costs of changes made

for reasons exogenous to the program itself. The *mistakes* bin is intended to capture the growth in the cost of obtaining the program approved at MS II. (Note that by these criteria, the costs of changes in a system needed to meet MS II requirements should be classified as *mistakes*, not *decisions*.)

The categorization is based on details of cost growth reported in the SARs. The SARs report cost growth attributed by the program manager to each of 7 categories, defined as follows:⁵

1. Economic change—A change that is solely due to price-level changes
2. Quantity change—A cost variance that is due to a change in the number of units acquired
3. Schedule change—Costs resulting from a change in a procurement or delivery schedule, completion date, or intermediate milestone for development or production
4. Engineering change—Cost increases or decreases that are due to an alteration in the physical or functional characteristics of a system or item delivered
5. Estimating Change—Changes that are due solely to correction of previous estimating errors or refinements of a current estimate
6. Other—Cost variances that are due to unforeseen events or not covered in any other category (e.g., natural disaster, strike)
7. Support change—Any change in cost, regardless of the reason, associated with support equipment for the major item of hardware.

Over the life of a major system, several dozen to several hundred variances typically will be posted. Some of these are explained in the SAR text; others are posted in one of the SAR categories without specific explanation. PA&E had each of these examined and classified as the result of a *decision* or a *mistake*.⁶ The rule used in classifying the individual elements of cost growth was that anything not clearly a *decision* is counted as a *mistake*. Ordinarily, the classification was based only on information reported in the SAR. Thus, only cost variances described in the SARs in a

⁵ These definitions, except for two minor differences, are those provided by Hough, "Pitfalls in Calculating Cost Growth from Selected Acquisition Reports," pp. 5–6.

⁶ The classifications were done by the Naval Sea Systems Support Office (NAVSHIPSO) under contract with PA&E and in accordance with PA&E guidelines. PA&E cost analysts also reviewed the NAVSHIPSO classifications.

way that clearly identified them as exogenous changes in the content approved at MS II would be included in *decisions*.

Table A-1 reports the *decisions* and *mistakes* components of development cost growth. The average *mistakes* component of cost growth is 18 percent for procurement; the estimated decisions component is 10 percent.

**Table A-1. Average Growth in MDAP Cost from MS II
Due to *Decisions* and *Mistakes***

	Percentage Growth	
	Development	Procurement
<i>Decisions</i>	21	10
<i>Mistakes</i>	24	18
Total	45	28

Source: Preliminary DoD data.

APPENDIX B: SYSTEMS INCLUDED IN SAMPLE

Table B-1. Systems in Our Usable Sample

System Name	MSII Year	Procurement Cost ^a
LGM-118A Peacekeeper Rail Garrison	1988	3,530
FGM-148A Javelin Advanced Anti-Tank Weapon System-Medium (AAWS-M)	1989	3,442
Advanced Field Artillery Tactical Data System (AFATDS)	1990	688
F-22 Advanced Tactical Fighter (ATF)	1991	52,271
F/A-18 E/F Super Hornet	1992	57,965
SYQ-23 Joint Service Imagery Processing System (JSIPS)	1992	362
LGM-30 Minuteman III Guidance Replacement Program (GRP)	1993	1,030
T-AKR 295 Strategic Sealift	1993	5,445
AV-8B Harrier Remanufacture	1994	2,089
E-2C Hawkeye Airborne Early Warning (AEW)	1994	2,826
M2/M3 Bradley FVS Upgrade	1994	3,079
MIM-104 Patriot Advanced Capability 3 (PAC-3)	1994	3,316
RGM-109 Tomahawk Baseline Improvement Program (T-BIP)	1994	1,306
V-22 Osprey (U.S. Navy)	1994	25,919
ALQ-212(V) Advanced Threat Infrared Countermeasure/Common Missile Warning System (ATIRCM/CMWS)	1995	2,247
BLU-108 Joint Standoff Weapon (JSOW) Unitary	1995	3,848
Cooperative Engagement Capability (CEC)	1995	1,234
Joint Direct Attack Munitions (JDAM)	1995	1,725
National Airspace System (NAS)	1995	588
SSN 774 Virginia Class New Attack Submarine	1995	47,202
AIM-9X Sidewinder Missile	1996	2,014
LPD-17 Amphibious Transport Dock Ship	1996	9,373
Navy Area Theater Ballistic Missile Defense (TBMD)	1997	3,459
Evolved Expendable Launch Vehicle (EELV)	1998	N/A
Joint Air-to-Surface Standoff Missile (JASSM)	1998	N/A

^a MS II procurement cost in millions of FY2000 dollars from the PA&E database.

ABBREVIATIONS

ADM	Acquisition Decision Memorandum
APB	Acquisition Program Baseline
AUPC	average unit procurement cost
BCP	Budget Change Proposal
CAIG	Cost Analysis Improvement Group
CARD	Cost Analysis Requirements Description
DAB	Defense Acquisition Board
DoD	Department of Defense
EMD	Engineering and Manufacturing Development
FY	fiscal year
FYDP	Future Years Defense Program
ICE	Independent Cost Estimate
IDA	Institute for Defense Analyses
MDA	Milestone Decision Authority
MDAP	major defense acquisition program
MS	Milestone
OSD	Office of the Secretary of Defense
PA&E	Program Analysis and Evaluation
PAUC	program acquisition unit cost
PCP	Program Change Proposal
PM	program manager
SAR	Selected Acquisition Report
SCP	Service Cost Position
SDD	System Development and Demonstration
USD(AT&L)	Under Secretary of Defense (Acquisition, Technology and Logistics)

REPORT DOCUMENTATION PAGE*Form Approved*
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 08-2005			2. REPORT TYPE Final		3. DATES COVERED (From - To) Jul 2004-Aug 2005	
4. TITLE AND SUBTITLE The Accuracy of Independent Estimates of the Procurement Costs of Major Systems					5a. CONTRACT NUMBER DASW01 04 C 0003	
					5b. GRANT NUMBER	
					5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S), Last Names First McNicol, David L. Tyson, Karen W. Hiller, John R. Cloud, Harley A. Minix, Joshua A.					5d. PROJECT NUMBER	
					5e. TASK NUMBER AO-7-2454	
					5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882					8. PERFORMING ORGANIZATION REPORT NUMBER IDA Paper P-3989	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of the Director, Acquisition Resources and Analysis Room 3D161 3020 Defense Pentagon, Room 3D161 Washington, DC 20301					10. SPONSOR/MONITOR'S ACRONYM(S) OUSD(AT&L)/ARA	
					11. SPONSORING/MONITORING AGENCY REPORT NUMBER	
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT This paper reports the results of an assessment of the accuracy of the independent estimates of procurement costs the Department of Defense uses in milestone reviews of major acquisition programs. IDA initially examined 63 major programs that received authorization to begin Engineering and Manufacturing Development during the years 1985 through 1998. The data needed to evaluate the accuracy of the independent procurement cost estimates could be located for only 25 of these programs. IDA found that the independent cost estimate was reasonably accurate for 18 of those 25 cases. In the remaining cases, the independent estimate was sometimes substantially too high and sometimes substantially too low.						
15. SUBJECT TERMS Cost Estimates; Military Procurement; Department of Defense; Accuracy						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report	18. NUMBER OF PAGES 48	19a. NAME OF RESPONSIBLE PERSON (Last Name First) Irby, Wendell	
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) (703) 695-9692	